

On Effectiveness of Backlog Bounds Using Stochastic Network Calculus in 802.11

Yue Wang (corresponding author)

School of Information

Central University of Finance and Economics

Address: 39 South College Road, Haidian District, Beijing, China 100081

Abstract

Network calculus is a powerful methodology of characterizing queueing processes and has wide applications, but few works on applying it to 802.11 by far. In this paper, we take one of the first steps to analyze the backlog bounds of an 802.11 wireless LAN using stochastic network calculus. In particular, we want to address its effectiveness on bounding backlogs. We model a wireless node as a single server with impairment service based on two best-known models in stochastic network calculus: Jiang's and Ciucu's. Interestingly, we find that the two models can derive equivalent stochastic service curves and backlog bounds in our studied case. We prove that the network-calculus bounds imply stable backlogs as long as the average rate of traffic arrival is less than that of service, indicating the theoretical effectiveness of stochastic network calculus in bounding backlogs. From A. Kumar's 802.11 model, we derive the concrete stochastic service curve of an 802.11 node and its backlog bounds. We compare the derived bounds with ns-2 simulations and find that the former are very loose and we discuss the reasons. And we show that the martingale and independent case analysis techniques can improve the bounds significantly. Our work offers a good reference to applying stochastic network calculus to practical scenarios.

Key words: stochastic network calculus, Backlog, 802.11

1 Introduction

Network calculus provides an elegant way to characterize traffic and service processes of network and communication systems. Unlike traditional queueing theory in which one has to make strong assumptions on arrival or service processes (e.g., Poisson arrival process, exponential service distribution, etc) so as to derive closed-form solutions in queueing networks [1], network calculus allows general

arrival and service processes. Instead of getting exact solutions, one derives network backlog and delay bounds by network calculus. Deterministic network calculus is mature in theory [2] [3] [6] [7]. However, most traffic and service processes are stochastic and deterministic network calculus is often not applicable to them. Therefore, stochastic network calculus was proposed to deal with stochastic arrival and service processes [7]- [18] [25].

Numerous applications of it have been found in communication networks and even in management science, and we cite some of them [19]- [27]. However, few works have been made on applying it to multiple access communication networks such as 802.11 Wireless LANs [26] [25]. In the paper, we take one of the first steps to apply stochastic network calculus to an 802.11 wireless LAN (WLAN). In particular, we want to address the effectiveness of stochastic network calculus on bounding backlogs in 802.11, with the following sub-problems:

- Under what condition can we derive stable backlogs using network calculus?
- How to derive the *concrete* stochastic service curve an 802.11 node?
- Are the derived backlog bounds tight compared with ns-2 simulations? And how to improve them?

We model a wireless node as a single server with impairment service based on two best-known models in stochastic network calculus: Jiang's [16] and Ciucu's [17]. And we make the following *new* contributions on this topic:

- We compare Jiang's and Ciucu's model and find that they can derive equivalent stochastic service curves and backlog bounds in our studied case.
- We prove that the network-calculus backlog bounds imply stable backlog as long as the average rate of traffic arrival is less than that of service, indicating that stochastic network calculus is effective in bounding backlogs *theoretically*.
- From A. Kumar's 802.11 model, we derive the *concrete* stochastic service curve of an 802.11 node [30] and give the numerical computation methods. From the service curve we then derive backlog bounds.
- We observe the derived bounds are loose when compared with ns-2 simulations. However, the martingale and independent case analysis techniques can improve the bounds significantly.

Note that when we prove a statement in this paper, we call it *Propositions* to differentiate the existing theorems in the literature (see Proposition 1-3).

This paper is organized as follows. In Section 2, we give a brief overview of stochastic network calculus. In particular, we present the classic models of Jiang's and Ciucu's and we also discuss the martingale and independent case analysis techniques. In Section 3, we present the network calculus model of a wireless node based on Jiang's and Ciucu's model. We compare the two models and find that they are equivalent in deriving stochastic service curves and backlog bounds in our studied case. We also prove the stability condition by the theory of stochastic network

calculus in this section. In Section 4, we derive the backlog bounds of an 802.11 node and the critical part is to derive its *concrete* stochastic service curve. In Section 5, we compare the derived backlog bounds with ns-2 simulation results under Poisson traffic arrivals. In particular, we show that the martingale and independent case analysis techniques can improve the bounds significantly. In Section 6, we give related works and highlight our contributions. Finally, Section 7 concludes the paper and points out some future works.

2 Stochastic Network Calculus

In this section, we first review basic terms of network calculus and then cite some results of the stochastic network calculus theory used in our paper. Jiang classified stochastic arrival curves as the types of *ta* (*traffic amount centric*), *vb* (*virtual backlog centric*) and *mb* (*max virtual backlog centric*), and classified stochastic service curve as *ws* (*weak stochastic*) and *sc* (*stochastic*). In this paper, we adopt *ta* and *mb* arrival curves and the *ws* service curve, as currently they provides tightest backlog bounds¹. Note that we just say "*stochastic service curve*" in our paper which means the *ws* one.

2.1 Basic Terms of Network Calculus

We consider a discrete time system where time is slotted ($t = 0, 1, 2, \dots$). A process is a function of time t . By default, we use $A(t)$ to denote the *arrival process* to a network element with $A(0) = 0$. $A(t)$ is the total amount of traffic arrived to this network element up to time t . We use $A^*(t)$ to denote the *departure process* of the network element with $A^*(0) = 0$. $A^*(t)$ is the total amount of traffic departed from the network element up to time t . Let \mathcal{F} ($\bar{\mathcal{F}}$) represents the set of non-negative wide-sense increasing (decreasing) functions. Clearly, $A(t) \in \mathcal{F}$ and $A^*(t) \in \bar{\mathcal{F}}$. For any process, say $A(t)$, we define $A(s, t) \equiv A(t) - A(s)$, for $s \leq t$. We define the backlog of the network element at time t by

$$B(t) = A(t) - A^*(t), \quad (1)$$

and the delay of the network element at t by

¹ As recently known by the network calculus community the bounding probability of a *mb* arrival curve is 1 for linear arrival curve functions, making its usage restrictive. So does the *sc* service curve as it is often derived from an impairment process with the *mb* arrival curve.

$$D(t) = \inf\{d : A(t) \leq A^*(t + d)\}. \quad (2)$$

Fig. 1 illustrates an example of $A(t)$ and $A^*(t)$ with $B(t)$ and $D(t)$ at $t = 10$.

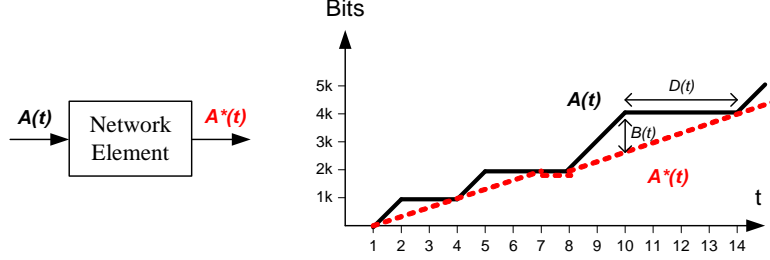


Fig. 1. Illustration of $A(t)$, $A^*(t)$, $B(t)$ and $D(t)$

In deterministic network calculus, $A(t)$ can be upper-bounded by an arrival curve. That is, for all $0 \leq s \leq t$, we have

$$A(s, t) \leq \alpha(t - s),$$

where $\alpha(t)$ is called the *arrival curve* of $A(t)$.

We say, *busy period* is a time period during which the backlog in the network element is always nonzero. For any busy period $(t_0, t]$, suppose we have

$$A^*(t) - A^*(t_0) \geq \beta(t - t_0),$$

if the network element provides a guaranteed service lower-bounded by $\beta(t - t_0)$ during the busy period. We can let t_0 be the beginning of the busy period, that is, the backlog at t_0 is zero or $A^*(t_0) = A(t_0)$. Therefore,

$$A^*(t) - A(t_0) \geq \beta(t - t_0).$$

The above equation infers $A^*(t) \geq \inf_{0 \leq s \leq t} [A(s) + \beta(t - s)]$, which can be written as

$$A^*(t) \geq A \otimes \beta(t), \quad (3)$$

where \otimes is called the operator of *min-plus convolution* and $\beta(t)$ is called the *service curve* of the network element.

2.2 Stochastic Network Calculus Theory

We consider a server S (i.e. the network element) fed with a flow A . In practice, A 's traffic and S 's service are often stochastic, which can not be hard bounded by some curves. That is, they can violate the curves but with certain probabilities (we call it *bounding function* here). The theory of stochastic network calculus can get probabilistic bounds for backlogs and delays of the server, suppose we can characterize A by a stochastic arrival curve and S by a stochastic service curve.

In this section, we just consider the derivation of backlog bounds as delay bounds are quite similar to the former. We first give some definitions. Then we cite some results in Jiang's and Ciucu's models [16] [17] and the construction. Lastly, we make a brief discussion on them.

2.2.1 Definitions

Definition 1 (ta stochastic arrival curve) A flow is said to have a ta (traffic-amount-centric) stochastic arrival curve $\alpha \in \mathcal{F}$ with bounding function $f \in \bar{\mathcal{F}}$, denoted by $A \sim_{ta} \langle f, \alpha \rangle$, if for all $s, t \geq 0$ ($s \leq t$) and all $x \geq 0$, there holds

$$P\{A(s, t) - \alpha(t - s) > x\} \leq f(x). \quad (4)$$

Definition 2 (vb stochastic arrival curve) A flow is said to have a vb (virtual-backlog-centric) stochastic arrival curve $\alpha \in \mathcal{F}$ with bounding function $f \in \bar{\mathcal{F}}$, denoted by $A \sim_{vb} \langle f, \alpha \rangle$, if for all $t \geq 0$ and all $x \geq 0$, there holds

$$P\left\{\sup_{0 \leq s \leq t} [A(s, t) - \alpha(t - s)] > x\right\} \leq f(x). \quad (5)$$

We can see that $A \sim_{vb} \langle f, \alpha \rangle$ implies $A \sim_{ta} \langle f, \alpha \rangle$, since $P\{A(s, t) - \alpha(t - s) > x\} \leq P\{\sup_{0 \leq s \leq t} [A(s, t) - \alpha(t - s)] > x\}$.

Definition 3 (Stochastic Service Curve) A server S is said to provide a (weak) stochastic service curve $\beta \in \mathcal{F}$ with bounding function $g \in \bar{\mathcal{F}}$, denoted by $S \sim_{ws} \langle g, \beta \rangle$ (or just $S \sim \langle g, \beta \rangle$), if for all $t \geq 0$ and all $x \geq 0$, there holds

$$P\{A \otimes \beta(t) - A^*(t) > x\} \leq g(x). \quad (6)$$

Definition 4 (Leftover Service) Consider a server S provides the ideal service curve $\hat{\beta}(t)$ with the impairment process I to a flow. Then, during any backlogged

period $(s, t]$, the output flow $A^*(s, t)$ from the server satisfies

$$A^*(s, t) \geq \hat{\beta}(t - s) - I(s, t). \quad (7)$$

$\hat{\beta}(t) - I(t)$ is the leftover service received by the given flow.

The definition of leftover service (also called *stochastic strict server* in [16]) can be applied to many scenarios such as cross traffic and wireless channels.

Definition 5 (θ -MER) A process A 's minimum envelope rate with respect to θ (θ -MER), denoted by $\rho^*(\theta)$, is defined as follows:

$$\rho^*(\theta) = \limsup_{t \rightarrow \infty} \frac{1}{\theta t} \sup_{s \geq 0} \log Ee^{\theta A(s, s+t)}. \quad (8)$$

We say that A has an envelope rate with respect to θ (θ -ER), denoted by $\rho(\theta)$, if $\rho(\theta) \geq \rho^*(\theta)$.

Definition 6 ($(\sigma(\theta), \rho(\theta))$ -upper constrained) A process A is said to be $(\sigma(\theta), \rho(\theta))$ -upper constrained for some $\theta > 0$, if for all $0 \leq s \leq t$, we have

$$\frac{1}{\theta} \log Ee^{\theta A(s, t)} \leq \rho(\theta)(t - s) + \sigma(\theta). \quad (9)$$

We can derive stochastic arrival and service curves from the $(\sigma(\theta), \rho(\theta))$ -upper constrained characterization (Section 2.3).

Definition 7 (Average Rate) The average rate of a process A , denoted by a_A , is defined as

$$a_A = \limsup_{t \rightarrow \infty} \sup_{s \geq 0} \frac{EA(s, s+t)}{t}. \quad (10)$$

Definition 8 (Stable Backlog Bound) The backlog $B(t)$ is stable, if for all t ,

$$EB(t) < C < \infty, \quad (11)$$

where C is a finite constant value. We say that the backlog bounds are stable if they can derive stable backlogs.

2.2.2 Jiang's Model

Jiang's model deals with *vb arrival curves* and stochastic service curves. We have the following theorems for leftover service curves and backlog bounds.

Theorem 1 (Jiang's Leftover Stochastic Service Curve) *Suppose a server S providing the ideal service curve $\hat{\beta}(t)$ with the impairment process I . If I has a vb stochastic arrival curve, i.e., $I \sim_{vb} \langle g, \gamma \rangle$, then the server provides the flow the leftover stochastic service curve $S \sim \langle g, \beta \rangle$ and*

$$\beta(t) = \hat{\beta}(t) - \gamma(t). \quad (12)$$

Theorem 2 (Jiang's Backlog Bounds) *If the flow A has a vb stochastic arrival curve $A \sim_{vb} \langle f, \alpha \rangle$ and the server S provides a stochastic service curve $S \sim \langle g, \beta \rangle$ to the flow, then the backlog $B(t)$ of the flow in the server at time t satisfies:*

$$P\{B(t) > x + \sup_{s \geq 0} [\alpha(s) - \beta(s)]\} \leq f \otimes g(x), \quad (13)$$

for all $t \geq 0$ and all $x \geq 0$.

2.2.3 Ciucu's Model

Ciucu's model deals with *ta arrival curves* and stochastic service curves.

In fact, we can derive vb arrival curves from ta arrival curves by introducing the function $\delta(t) = \delta \cdot t$ (δ is an adjustable constant). The following lemma states this.

Lemma 1 (ta to vb Arrival Curves) *Suppose A is a ta stochastic arrival curve, $A \sim_{ta} \langle f, \alpha \rangle$, then $A \sim_{vb} \langle \tilde{f}, \alpha_\delta \rangle$ with $\alpha_\delta(t) \equiv \alpha(t) + \delta t$ and its bounding function $\tilde{f}(x, \delta) = \sum_{k=0}^{\infty} f(x + k\delta)$ (suppose the sum is finite).*

The derivations are as follows.

$$\begin{aligned}
& P\left\{\sup_{0 \leq s \leq t} [A(s, t) - \alpha_\delta(t - s)] > x\right\} \\
& \leq \sum_{s=0}^t P\{A(s, t) - \alpha_\delta(t - s) > x\} \\
& = \sum_{s=0}^t P\{A(s, t) - \alpha(t - s) > x + \delta(t - s)\} \\
& \leq \sum_{s=0}^t f(x + \delta(t - s)) \leq \sum_{k=0}^{\infty} f(x + k\delta).
\end{aligned} \tag{14}$$

We have the following theorems for leftover service curves and backlog bounds in Ciucu's model. Actually, we can derive these results by first converting ta arrival curves to vb ones and then applying Jiang's theorems.

Theorem 3 (Ciucu's Leftover Stochastic Service Curve) *Suppose a server S providing the ideal service curve $\hat{\beta}(t)$ with the impairment process I . If I has a ta stochastic arrival curve, i.e., $I \sim_{ta} \langle g, \gamma \rangle$, then the server provides the flow the leftover stochastic service curve $S \sim \langle \tilde{g}, \beta \rangle$ and*

$$\beta(t) = \hat{\beta}(t) - \gamma_\delta(t), \tag{15}$$

where $\gamma_\delta(t) \equiv \gamma(t) + \delta t$ and $\tilde{g}(x, \delta) \equiv \sum_{k=0}^{\infty} g(x + k\delta)$ by definition.

Theorem 4 (Ciucu's Backlog Bounds) *If the flow A has a ta stochastic arrival curve $A \sim_{ta} \langle f, \alpha \rangle$ and the server S provides a stochastic service curve $S \sim \langle g, \beta \rangle$ to the flow, then the backlog $B(t)$ of the flow in the server satisfies: for all $t \geq 0$ and all $x \geq 0$,*

$$P\{B(t) > x + \sup_{s \geq 0} [\alpha_\delta(s) - \beta(s)]\} \leq \tilde{f} \otimes g(x), \tag{16}$$

where $\alpha_\delta(t) \equiv \alpha(t) + \delta t$ and $\tilde{f}(x, \delta) \equiv \sum_{k=0}^{\infty} f(x + k\delta)$ by definition.

Note that Ciucu's can deal with ta arrival curves while Jiang's can not, by introducing a $\delta > 0$ to trades smaller service for finite bounding functions.

2.3 Computation of Stochastic Arrival/Service Curves

We will show in this subsection how to calculate stochastic arrival and service curves from the $(\sigma(\theta), \rho(\theta))$ -upper constrained characterization [7].

Theorem 5 (Arrival Curves of $(\sigma(\theta), \rho(\theta))$ -upper constrained) Suppose $A(t)$ is $(\sigma(\theta), \rho(\theta))$ -upper constrained, then it has a ta stochastic arrival curve $A \sim_{ta} < f, \alpha >$, where

$$\begin{aligned}\alpha(t) &= r \cdot t \\ f(x) &= e^{\theta\sigma(\theta)} \cdot e^{-\theta x},\end{aligned}\tag{17}$$

for any $r \geq \rho(\theta)$ and $x \geq 0$. And A has a vb stochastic arrival curve $A \sim_{vb} < f, \alpha >$, where

$$\begin{aligned}\alpha(t) &= r \cdot t \\ f(x) &= \frac{e^{\theta\sigma(\theta)}}{1 - e^{\theta(\rho(\theta)-r)}} \cdot e^{-\theta x},\end{aligned}\tag{18}$$

for any $r > \rho(\theta)$ and $x \geq 0$.

Note that we have $r \geq \rho(\theta)$ in ta and $r > \rho(\theta)$ in vb. And Eq.(18) applies Boole's inequality to the bound functions $f(x)$ which are loose in general.

How to derive stochastic service curves? If we can model the server S with the ideal service curve $\hat{\beta}$ with the impairment process $I(t)$, we can first characterize $I(t)$ by vb (ta) arrival curves, and then we use Theorem 1 (Theorem 3) to get its stochastic service curves.

The following theorem states the relation between θ -ER and $(\sigma(\theta), \rho(\theta))$ -upper constrained. We will use it in proving the stability condition of backlog bounds in Section 3.4.

Theorem 6 (θ -ER vs $(\sigma(\theta), \rho(\theta))$ -upper constrained) If the process $A(t)$ has a θ -envelop rate (θ -ER) $\rho(\theta) < \infty$, then for every $\epsilon > 0$ there exists $\sigma_\epsilon(\theta) < \infty$ so that A is $(\sigma_\epsilon(\theta), \rho(\theta) + \epsilon)$ -upper constrained.

2.4 Improvement on Bounds

There are two ways of improving bounds in current literature. One way is to apply independent case analysis. The other way is to improve the bounding functions of stochastic service curves for time-independent arrivals.

The first way says that suppose the impairment process of the server S is independent from the traffic arrival process, we can derive tighter backlog bounds using independent probability analysis.

Theorem 7 (Backlog Bounds under Independent Cases) *Suppose the server S provides the flow (satisfying $A \sim_{vb} \langle f, \alpha \rangle$) the ideal service curve $\hat{\beta}(t)$ with the impairment process $I \sim_{vb} \langle g, \gamma \rangle$ (thus $S \sim \langle g, \beta \rangle$ where $\beta(t) = \hat{\beta}(t) - \gamma(t)$). Suppose A and I are independent, we have*

$$\begin{aligned} & P\{B(t) > \sup_{s \geq 0} [\alpha(s) - \beta(s)] + x\} \\ & \leq \sum_{k=0}^x (\bar{g}(k) - \bar{g}(k-1)) \bar{f}(x-k) \end{aligned} \quad (19)$$

where $\bar{f}(x) = 1 - f(x)$, $\bar{g}(x) = 1 - g(x)$, and we set $\bar{g}(-1) = 0$.

This theorem of independent case analysis can be applied to Ciucu's model. However, we first need to convert the ta arrival curves to the vb ones with new bounding functions $\hat{f}(x, \delta)$ and $\hat{g}(x, \delta)$ by Lemma 1. Then we apply the above theorem by plugging in \hat{f} and \hat{g} .

Another way of tightening backlog bounds is to derive tighter bounding functions of stochastic arrival and service curves. Ciucu first proposed to use martingale to tighten the bounds for $M/M/1$ and $M/D/1$ queues [18]. In the following proposition, we provide a more general result following his idea. The proof is given in Appendix-A.

Proposition 1 (vb Arrival Curves of Time-Independent Process) *Suppose $A(t)$ is $(\sigma(\theta), \rho(\theta))$ -upper constrained. On condition that $a(t) \equiv A(t) - A(t-1)$ is independent of each t , it has a vb stochastic arrival curve $A \sim_{vb} \langle f, \alpha \rangle$, where*

$$\begin{aligned} \alpha(t) &= r \cdot t \\ f(x) &= e^{-\theta x}, \end{aligned} \quad (20)$$

for any $r \geq \rho(\theta) + \sigma(\theta)$ and $x \geq 0$.

2.5 Discussion on Jiang's and Ciucu's Models

The key difference between Jiang's and Ciucu's models is: Jiang use vb traffic arrival curves while Ciucu uses ta ones variant stochastic service curves. Which one can derive tighter backlog bounds?

In general, ta arrival curves provide tighter bounding functions than vb. Actually, $A \sim_{vb} \langle f, \alpha \rangle$ implies $A \sim_{ta} \langle f, \alpha \rangle$ and the inverse is not true generally. In

particular, the bounding function of τ_a is tighter than that of τ_b (especially when r is close to $\rho(\theta)$) in Theorem 5. But one can not conclude that Ciucu's model is always better than Jiang's, as it has looser bounding functions for leftover service curves and backlog bounds (see Theorem 3 and Theorem 4). The situation becomes even more uncertain when consider time-independent processes and independent A and I . Interestingly, we find that two models can derive equivalent stochastic service curves and backlog bounds in our studied case (Section 3.3).

3 A Wireless Node's Network Calculus Model

In this section, we model a general wireless node by stochastic network calculus. In general, we can define one time slot ($t = 1$) to be any small duration and measure traffic amount in any unit (e.g. bits, bytes or packets).

We consider a wireless node. Let $A(t)$ denote the traffic arrived at the node from the application layer. We assume A is $(\sigma_A(\theta_1), \rho_A(\theta_1))$ -upper constrained, which is a right assumption for many cases.

We model the service of a wireless node as an ideal server curve with an impairment process. In fact, the channel is shared by the other node in a WLAN and transmission errors occur due to path loss, fading and collisions, which contribute to the impairment process. Let the channel capacity be c traffic units per slot. The departure process $A^*(s, t) = \hat{\beta}(s, t) - I(s, t)$ during any backlogged period $[s, t]$, where $\hat{\beta}(t) = c \cdot t$ is the ideal service curve and I is the impairment process. Since $I(s, t) \leq c \cdot (t - s)$, there exist $\sigma_I(\theta_2)$ and $\rho_I(\theta_2)$ so that I is $(\sigma_I(\theta_2), \rho_I(\theta_2))$ -upper constrained.

In here, θ_1 and θ_2 are adjustable parameters. We will show Section 4 how to calculate $\rho_A(\theta_1)$, $\sigma_A(\theta_1)$, $\rho_I(\theta_2)$ and $\sigma_I(\theta_2)$ for an 802.11 node.

3.1 Jiang's Backlog Bounds

Because A is $(\sigma_A(\theta_1), \rho_A(\theta_1))$ -upper constrained, by Theorem 5, $A \sim_{vb} < f, \alpha >$ where

$$\begin{aligned} \alpha(t) &= r_A \cdot t \\ f(x) &= \frac{e^{\theta_1 \sigma_A(\theta_1)}}{1 - e^{\theta_1 (\rho_A(\theta_1) - r_A)}} \cdot e^{-\theta_1 x}, \end{aligned} \quad (21)$$

for any $r_A > \rho_A(\theta_1)$.

In the same way, $I \sim_{vb} < g, \gamma >$ where

$$\begin{aligned}\gamma(t) &= r_I \cdot t \\ g(x) &= \frac{e^{\theta_2 \sigma_I(\theta_2)}}{1 - e^{\theta_2(\rho_I(\theta_2) - r_I)}} \cdot e^{-\theta_2 x},\end{aligned}\tag{22}$$

for any $r_I > \rho_I(\theta_2)$.

By Theorem 1, the node provides a stochastic service curve $S \sim < g, \beta >$, where

$$\beta(t) = (c - r_I) \cdot t,\tag{23}$$

for any $c > r_I$.

Finally, by Theorem 2, we must let $\alpha(t) \leq \beta(t)$, i.e., $r_A \leq c - r_I$, in order to get meaningful backlog bounds which are $P\{B(t) > x\} \leq f \otimes g(x)$.

We note that $f(x)$ ($g(x)$) is the decreasing function of r_A (r_I). Considering the above conditions, we get the following optimal backlog bounds,

$$\begin{aligned}P\{B(t) > x\} &\leq \min_{\theta_1, \theta_2, r_A, r_I} [f \otimes g(x)] \\ \text{subject to} \\ r_A &> \rho_A(\theta_1), r_I > \rho_I(\theta_2) \\ r_A + r_I &= c \\ \theta_1, \theta_2 &> 0.\end{aligned}\tag{24}$$

In here, $\rho_A(\theta_1)$ ($\rho_I(\theta_2)$) is the function of θ_1 (θ_2).

3.2 Ciucu's Backlog Bounds

Because A is $(\sigma_A(\theta_1), \rho_A(\theta_1))$ -upper constrained, by Theorem 5, $A \sim_{ta} < f, \alpha >$, where

$$\begin{aligned}\alpha(t) &= r_A \cdot t \\ f(x) &= e^{\theta_1 \sigma_A(\theta_1)} \cdot e^{-\theta_1 x},\end{aligned}\tag{25}$$

for any $r_A \geq \rho_A(\theta_1)$.

In the same way, $I \sim_{ta} < g, \gamma >$, where

$$\begin{aligned}\gamma(t) &= r_I \cdot t \\ g(x) &= e^{\theta_2 \sigma_I(\theta_2)} \cdot e^{-\theta_2 x},\end{aligned}\tag{26}$$

for any $r_I \geq \rho_I(\theta_2)$.

By Lemma 1, we have $A \sim_{vb} < \tilde{f}, \alpha_{\delta_1} >$ where

$$\begin{aligned}\alpha_{\delta_1}(t) &= (r_A + \delta_1) \cdot t \\ \hat{f}(x, \delta_1) &= \frac{e^{\theta_1 \sigma_A(\theta_1)}}{1 - e^{-\theta_1 \delta_1}} \cdot e^{-\theta_1 x},\end{aligned}\tag{27}$$

for any $r_A \geq \rho_A(\theta_1)$ and $\delta_1 > 0$. In here, we can get the close form of $\tilde{f}(x, \delta) = \sum_{k=0}^{\infty} f(x + k\delta)$ for the particular $f(x)$ in Eq.(25).

In the same way, $I \sim_{vb} < \tilde{g}, \gamma_{\delta_2} >$ where

$$\begin{aligned}\gamma_{\delta_2}(t) &= (r_I + \delta_2) \cdot t \\ \hat{g}(x, \delta_2) &= \frac{e^{\theta_2 \sigma_I(\theta_2)}}{1 - e^{-\theta_2 \delta_2}} \cdot e^{-\theta_2 x},\end{aligned}\tag{28}$$

for any $r_I \geq \rho_I(\theta_2)$ and $\delta_2 > 0$.

By Theorem 3, the node provides a stochastic service curve $S \sim < \tilde{g}, \beta >$, where

$$\beta_{-\delta_2}(t) = (c - r_I - \delta_2) \cdot t,\tag{29}$$

for any $c > r_I + \delta_2$.

Finally, by Theorem 4, we must have $\alpha_{\delta_1}(t) \leq \beta_{-\delta_2}(t)$, i.e., $r_A + \delta_1 \leq c - r_I - \delta_2$ in order to get meaningful backlog bounds which are $P\{B(t) > x\} \leq \tilde{f} \otimes \tilde{g}(x)$. We note that $f(x)$ ($g(x)$) is the decreasing function of δ_1 (δ_2). Considering the above conditions, we get the following optimal backlog bounds,

$$\begin{aligned}P\{B(t) > x\} &\leq \min_{\theta_1, \theta_2, \delta_1, \delta_2, r_A, r_I} [\tilde{f} \otimes \tilde{g}(x)] \\ &\text{subject to} \\ r_A &> \rho_A(\theta_1), r_I > \rho_I(\theta_2) \\ r_A + r_I + \delta_1 + \delta_2 &= c \\ \theta_1, \theta_2 > 0, \delta_1, \delta_2 &> 0.\end{aligned}\tag{30}$$

In here, $\rho_A(\theta_1)$ ($\rho_I(\theta_2)$) is the function of θ_1 (θ_2).

3.3 Equivalent Bounds in Two Models

We find that the two models actually can derive the same stochastic service curves and backlog bounds. To understand this, we note that the key difference is the traffic model. The following proposition shows that we can derive the same vb arrival curves from the two models.

Proposition 2 (ta vs vb Arrival Curves) *If A is a $(\sigma(\theta), \rho(\theta))$ -upper constrained process, its vb arrival curve immediately generated by applying Theorem 5's Eq.(18) and the one generated by applying Theorem 5's Eq.(17) and then Lemma 1 are equivalent.*

Proof: Following the discussions above, the vb arrival curves generated immediately by applying Theorem 5's Eq.(18) are $A \sim_{vb} < f, \alpha >$ where

$$\begin{aligned}\alpha(t) &= r \cdot t \\ f(x) &= \frac{e^{\theta\sigma(\theta)}}{1 - e^{\theta(\rho(\theta)-r)}} \cdot e^{-\theta x},\end{aligned}\tag{31}$$

for any $r > \rho(\theta)$.

The vb arrival curves by applying Theorem 5's Eq.(17) and then Lemma 1 (converted by the ta arrival curves) are $A \sim_{vb} < \hat{f}, \alpha_\delta >$ where

$$\begin{aligned}\alpha_\delta(t) &= (r + \delta) \cdot t \\ \hat{f}(x, \delta) &= \frac{e^{\theta\sigma(\theta)}}{1 - e^{-\theta\delta}} \cdot e^{-\theta x},\end{aligned}\tag{32}$$

for any $r \geq \rho(\theta)$ and $\delta > 0$.

For the same value of $(r + \delta)$ in Eq.(32), we should maximize δ to get tighter $\hat{f}(x, \delta)$; in other words, we should minimize r and let it to be $\rho(\theta)$. In this optimized case we find that Eq.(31) and Eq.(32) are in the same form. This establishes the equivalence between them. ■

This result can be applied to the vb arrival curves of impairment processes. Since Ciucu's model can be derived from Jiang's for the single-server case (Section 2.2.3), the two models can derive the same backlog bounds in the following proposition.

Proposition 3 (Bounds Equivalence in Two Models) *Consider a single server S with an ideal service curve $\hat{\beta}$ and an impairment process I . Suppose the traffic arrival process A and the impairment process I are $(\sigma(\theta), \rho(\theta))$ -upper constrained for some θ respectively, then the vb arrival curves, the stochastic service curve and backlog bounds derived by Jiang's and Ciucu's model are equivalent.*

We omit the proof which is got immediately from Proposition 2.

Note: By equivalence, we do not mean it is general for all situations. Actually, Ciucu's model can be extended to multiple concatenated nodes while Jiang's can not, and they are different models. Even for the single-node case, we only prove the equivalence property for *linear* arrival curves in our studied case. And it is still an open problem for more general cases.

3.4 Stability Condition

One fundamental question we need to address is under what condition we can derive *stable* backlog bounds (i.e., $EB(t) < \infty$) by stochastic network calculus. The following proposition shows the stability condition.

Proposition 4 (Stability Condition) *Suppose there exist θ -MERs (θ -Minimum Envelop Rates) for the traffic arrival process A and the impairment process I of the wireless node for $0 < \theta < \hat{\theta}$ where $\hat{\theta}$ is some constant value, then stochastic network calculus can derive stable backlogs if*

$$a_A < c - a_I, \quad (33)$$

where c is the transmission rate of the ideal channel, a_A and a_I are the average rate of A and I defined in Definition 7, respectively.

Proof:

The proof consists of two phases. First, we show that $a_A < c - a_I$ can lead to $r_A \leq c - r_I$. Next, we show that if $r_A \leq c - r_I$ then stochastic network calculus can derive $EB(t)$ which is less than a finite value.

We adopt Jiang's model in Section 3.1 where A is the traffic arrival process and I is the impairment process of the server S , since the two models are equivalent in our studied case (see Proposition 3). We have shown that $P\{B(t) > x\} \leq f \otimes g(x)$ if $r_A \leq c - r_I$ holds.

From Eq.(21) and (22), we let $\epsilon_1 = r_A - \rho_A(\theta_1)$ and $\epsilon_2 = r_I - \rho_I(\theta_2)$ for $\theta_1, \theta_2 > 0$ and $\epsilon_1, \epsilon_2 > 0$. To simplify the arguments, we let $\theta_1 = \theta_2 = \theta$ and $\epsilon_1 = \epsilon_2 = \epsilon$.

Thus, $r_A \leq c - r_I$ holds if

$$\rho_A(\theta) \leq c - \rho_I(\theta) - 2\epsilon. \quad (34)$$

From Theorem 6, we can construct the $(\sigma(\theta), \rho(\theta))$ -upper constrained characterization by letting $\rho_A(\theta) = \rho_A^*(\theta) + \epsilon$ and $\rho_I(\theta) = \rho_I^*(\theta) + \epsilon$ for any $\epsilon > 0$, where $\rho_A^*(\theta_1)$ and $\rho_I^*(\theta_2)$ are θ -MERs of A and I , respectively. And Eq.(34) holds if

$$\rho_A^*(\theta) \leq c - \rho_I^*(\theta) - 4\epsilon. \quad (35)$$

Because $\rho_A^*(\theta)$ exists, applying Taylor's expansion,

$$\begin{aligned} \rho_A^*(\theta) &= \lim_{t \rightarrow \infty} \sup_{s \geq 0} \frac{1}{\theta t} \sup \log \mathbb{E} e^{\theta A(s, s+t)} \\ &= \lim_{t \rightarrow \infty} \sup_{s \geq 0} \frac{1}{\theta t} \sup \log \mathbb{E} (1 + \theta A(s, s+t) + O(\theta^2 A(s, s+t)^2)) \\ &= \lim_{t \rightarrow \infty} \sup_{s \geq 0} \frac{1}{\theta t} \sup \log (1 + \theta \mathbb{E} A(s, s+t) + O(\theta^2 A(s, s+t)^2)) \\ &= \lim_{t \rightarrow \infty} \sup_{s \geq 0} \frac{1}{\theta t} \sup [\theta \mathbb{E} A(s, s+t) + O(\theta^2 A(s, s+t)^2)]. \end{aligned}$$

Let θ go to 0,

$$\lim_{\theta \rightarrow 0} \rho_A^*(\theta) = \lim_{t \rightarrow \infty} \sup_{s \geq 0} \frac{\mathbb{E} A(s, s+t)}{t} = a_A. \quad (36)$$

Similarly,

$$\lim_{\theta \rightarrow 0} \rho_I^*(\theta) = a_I. \quad (37)$$

Therefore, there exists some $\theta < \hat{\theta}$ so that $\rho_A^*(\theta) \leq a_A + \epsilon$ and $\rho_I^*(\theta) \leq a_I + \epsilon$. So Eq. (35) holds if

$$a_A \leq c - a_I - 6\epsilon. \quad (38)$$

Since ϵ can be arbitrarily small, Eq.(38) holds if

$$a_A < c - a_I. \quad (39)$$

Following the above derivations backwards, we prove that $a_A < c - a_I$ leads to $r_A < c - r_I$.

Next, we prove that stochastic network calculus can derive $EB(t)$ which is less than a finite value if $r_A < c - r_I$.

Since $f(x)$ and $g(x)$ are exponentially decreasing functions according to Eq. (21) and Eq. (22), we can show that $EB(t)$ is upper-bounded by some finite constant value as follows. Note that $B(t)$ is a discrete value in practice (e.g., in bits or packets).

$$\begin{aligned}
EB(t) &= \sum_{k=0}^{\infty} P\{B(t) = k + 1\} \cdot (k + 1) \\
&< \sum_{k=0}^{\infty} P\{B(t) > k\} \cdot (k + 1) \\
&\leq \sum_{k=0}^{\infty} f \otimes g(k) \cdot (k + 1) \\
&\leq \sum_{k=0}^{\infty} (f(\lfloor \frac{k}{2} \rfloor) + g(\lceil \frac{k}{2} \rceil)) \cdot (k + 1) < \infty.
\end{aligned} \tag{40}$$

■

Remarks: Since the proof is based on the theory of stochastic network calculus, it indicates that we can get stable backlog bounds by stochastic network calculus on the condition that the average arrival rate is less than the average service rate. As this condition is very general, stochastic network calculus is effective in theory.

4 An 802.11 Node's Network Calculus Model

In this section, we derive the backlog bounds for an 802.11 node. And the key part is to derive its stochastic service curve. We use Jiang's model in this section since the two models are equivalent in our studied case (Proposition 3).

For simplicity, we assume n *identical* nodes send packets to an AP (access point) and they share the wireless channel. All nodes operate in Distributed Coordination Function (DCF) mode with RTS/CTS turned off [28]. We assume that transmission errors only happen due to packet collisions and two packets are collided if their transmissions overlap in time. Besides, we assume that all DATA packets are of the same size for simplicity. We use Scenario 1 for illustration.

Scenario 1:

10 nodes send packets to one AP in a WLAN

The payload of a DATA packet is 256 bytes

Fig. 2. Scenario 1 of a wireless LAN

4.1 802.11 DCF Protocol

A node with a DATA packet (or simply packet) to transmit first senses channel state. If the channel is idle for the time of DIFS (distributed interframe space), the node transmits. Otherwise, if the channel is busy during the DIFS, the node backs off, i.e., the node defers channel access by a random number of *idle slots* ranging from 0 to $CW - 1$ within a contention window (CW). When the backoff counter reaches zero and expires, the node can access the channel. During the backoff period, if the node senses the channel is busy, it freezes the backoff counter and the backoff process is resumed once the channel is idle for a duration of DIFS. To avoid channel capture, a node must wait a random backoff time between two consecutive new packet transmissions, even if the channel is sensed idle. Once the packet is received successfully, the receiver will return an ACK after the duration of SIFS (short inter-frame space). SIFS is shorter than an idle slot so that there are no collisions caused by DATA packets and ACKs.

802.11 uses the truncated exponential backoff technique to set its CW . In 802.11b, the initial CW is $CW_{min} = 32$. Each time a collision occurs, CW doubles its size, up to $CW_{max} = 1024$. When the packet is successfully transmitted, CW is reset to CW_{min} . The packet is dropped when it is retransmitted for 6 times and still not transmitted successfully. Fig. 3 shows the parameters of 802.11b used in our paper.

The duration of an ACK is the duration of PHY header plus that of ACK header transmitted at *basic rate*, i.e., $\frac{(24+14) \cdot 8}{10^6} = 304\mu s \approx 16 \text{ idle slots}$. The duration of a DATA packet is the duration of PHY header transmitted at *basic rate* plus that of a MAC header and its upper-layer payload transmitted at *data rate*. For example, suppose the upper-layer payload is 256 bytes, then the duration of an DATA packet is $\frac{24 \cdot 8}{10^6} + \frac{(28+256) \cdot 8}{11 \cdot 10^6} = 398.5\mu s \approx 20 \text{ idle slots}$.

4.2 802.11 Service Curve

Since equal-sized DATA packets are considered, we measure traffic, service and backlog amount in packets in our paper. We measure time duration (e.g. SIFS, DIFS, DATA and ACK) in the unit of idle slots and define that *one time slot of network calculus* ($t = 1$) is equal to L idle slots, where

Basic rate	1 Mbps
Data rate	11 Mbps
PHY header	24 bytes
ACK header	14 bytes
MAC header	28 bytes
SIFS	10 μs
DIFS	50 μs
Idle slot	20 μs
CW_{min}	32
CW_{max}	1024

Fig. 3. 802.11b parameters

$$L = (DIFS + DATA + SIFS + ACK) \text{ in idle slots.} \quad (41)$$

Ideally, an 802.11 node transmits 1 packets per time slot (L idle slots in length). Suppose the DATA payload is 256 bytes, $L = 3 + 16 + 20 = 39$ *idle slots*.

Sometimes in the paper, "*idle slot*" refers to the time period which equal to the length of an idle slot and it may not be idle. To avoid this confusion, we will use "*idle slot*" (italic) to denote that the "idle slot" is indeed idle.

An 802.11 node can be modeled as an ideal server (1 packet transmitted per time slot) with the impairment process I which is due to contention with the other nodes in a WLAN. In practice, it is difficult to calculate I accurately since I depends on the complex interactions of traffic arrival and DCF. In this section, we assume the saturated state and use A. Kumar's fixed-point model of 802.11 [30]. This model is justified to be very accurate in practice [31].

We assume that the system is working at the saturated state, that is, the backlog at each node is always nonzero. For a given node, let τ denote its transmission attempt probability per *idle slot* and let η denote the conditional collision probability when it transmits a packet. We assume η is constant and independent for each transmission. Intuitively, this assumption becomes more accurate when the number of nodes n increases. In [30], the authors derived two general formulas relating τ to η . The first one is

$$\tau = \frac{1 + \eta + \eta^2 + \dots + \eta^6}{b_0 + \eta b_1 + \eta^2 b_2 + \dots + \eta^6 b_6}. \quad (42)$$

This equation can be explained as follows. The numerator is the expected number of transmission attempts of a packet. The denominator is the expected total backoff duration (in idle slots) of a packet, where b_i is the mean backoff duration after the i th collision plus 1 (the 1 refers to the first idle slot of a packet transmission). In 802.11, $b_i = \frac{2^i \cdot CW_{min}}{2}$ where $0 \leq i \leq 6$. A packet suffering 6 consecutive collisions will be dropped from its buffer. In our calculations, we do not consider packet drops. Since the probability of packet drops is very small, this simplification relaxes the backlog bounds very slightly.

The independence assumption of η implies that each transmission sees the system at steady state. Therefore, each node transmits with the same probability τ . This yields

$$\eta = 1 - (1 - \tau)^{n-1}. \quad (43)$$

Combining Eq. (42) and Eq. (43), we can solve τ and η .

We introduce the following terms. The probability of no transmissions at an *idle slot* in the WLAN, denoted by P_{nt} , is $(1 - \tau)^n$. The probability of having at least one transmission at an *idle slot* in the WLAN, denoted by P_t , is $1 - P_{nt}$. The probability of a *given node* starting a successful transmission at an *idle slot*, denoted by P_s , is $\tau(1 - \eta)$.

Fig. 4 plots Eq. (42) in dashed line and Eq. (43) in solid line when $n = 10, 20$ and 100 . The intersecting points are the solutions to η and τ . It can be seen from the figure that η increases and τ decreases as n increases. Consequently, P_{nt} increases while P_s and P_t decreases as n increases. When we consider the saturated state of the system, we actually consider all nodes contending the channel which gives the *worst-case* analysis of the impairment process of a given node and thus conservative backlog bounds of it. However, we argue that it is necessary because one applies network calculus to deriving the worst-case bounds.

In order to characterize the impairment process I of the given 802.11 node, it is crucial to know its moment generating function. Specifically, we want to calculate

$$M_I(t) = \sup_{s \geq 0} [E e^{\theta I(s, s+t)}], \quad (44)$$

and then we can know its $(\sigma(\theta), \rho(\theta))$ -upper constrained characterization (we replace θ_2 by θ here to simplify explanations).

We calculate Eq.(44) for a given 802.11 node as follows. Consider the duration of t time slots (i.e., tL idle slots) from s to $s + t$. Since we take the sup, we can assume that there is always a transmission by the other nodes at the first time slot $[s, s + 1]$, which gives the conservative estimation of $I(s, s + t)$ for the given node. We can

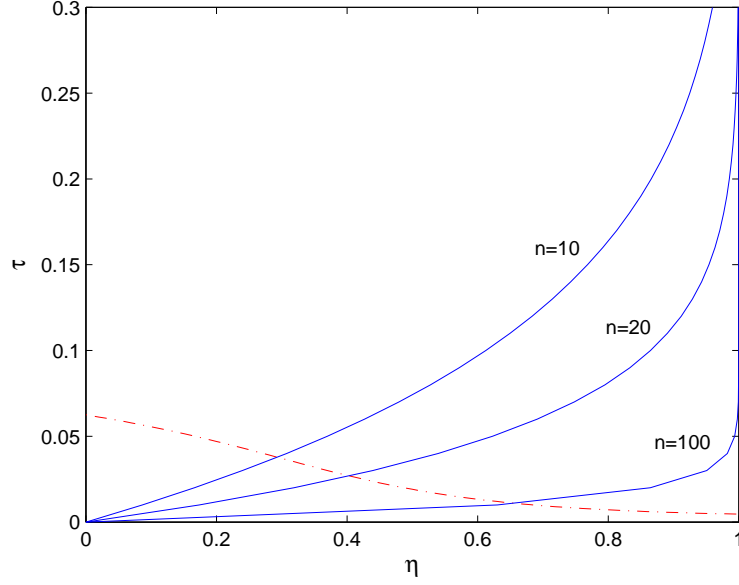


Fig. 4. The Plots of Eq. (42) and Eq. (43)

see that it is a good approximation as the transmissions by the other $n - 1$ nodes happen much frequently than the given node when n is large and also an *idle slot* is much smaller in length than L (in other words, the channel is often busy). We can see this point is right for Scenario 1 in the end of this subsection.

In the following, we consider probabilistic events in the remaining $t - 1$ time slots. There are two cases. *Case I*: The last transmission is "incomplete". *Case II*: Otherwise to Case I. By "incomplete", we means that the last transmission goes on for k idle slots ($1 \leq k \leq L - 1$) and get truncated due to the boundary of the last time slot. Let $\tilde{P}_s = P_s/P_t$ denote the condition probability of the given node's successful transmission on the condition that there is a transmission on the channel.

We first calculate $M_I(t)$ for case I. Suppose there are i complete transmissions and one incomplete transmission occupying k idle slots, its probability denoted by $p_{i,k}$ is $P_t \cdot C_{(t-i-1)L-k+i}^i P_t^i P_{nt}^{(t-i-1)L-k}$. In here, we use the fact that there are i complete transmissions, 1 incomplete transmission and thus $(t - i - 1)L - k$ idle slots in the remaining $t - 1$ time slots. Suppose there are j successful transmissions from the given node in the i complete transmissions, its probability denoted by $q_{j,i}$ is $C_i^j (\tilde{P}_s)^j (1 - \tilde{P}_s)^{i-j}$. When the last incomplete transmission is from the other nodes, $M_I(t)$ is $e^{\theta(t-j)}$; otherwise, the transmission is from the given node itself, $M_I(t)$ is $e^{\theta(t-j-k/L)}$.

Numerating all possible k , i and j , $M_I(t)$ under case I is

$$\begin{aligned}
& \sum_{k=1}^{L-1} \sum_{i=0}^{t-2} \sum_{j=0}^i \left(p_{i,k} q_{j,i} \tilde{P}_s e^{\theta(t-j-k/L)} + p_{i,k} q_{j,i} (1 - \tilde{P}_s) e^{\theta(t-j)} \right) \\
&= \sum_{k=1}^{L-1} \sum_{i=0}^{t-2} p_{i,k} (\tilde{P}_s e^{-\theta k/L} + 1 - \tilde{P}_s) (\tilde{P}_s e^{-\theta} + 1 - \tilde{P}_s)^i e^{\theta t}
\end{aligned} \tag{45}$$

Then we calculate $M_I(t)$ for case II. Let $p_i = C_{(t-i-1)L+i}^i P_t^i P_{nt}^{(t-i-1)L}$. Following the similar arguments as above, $M_I(t)$ for case II is:

$$\sum_{i=0}^{t-1} \sum_{j=0}^i p_i q_{j,i} e^{\theta(t-j)} = \sum_{i=0}^{t-1} p_i (\tilde{P}_s e^{-\theta} + 1 - \tilde{P}_s)^i e^{\theta t}. \tag{46}$$

Adding Eq.(45) and Eq.(46), finally we get $M_I(t)$.

In general, we do not have the analytical form of $M_I(t)$, so we resort to numerical methods to obtain $\sigma_I(\theta)$ and $\rho_I(\theta)$ (see Algorithm 1 in Appendix B). The algorithm is immediately inspired from Definition 6. Then we can use Eq. (22) and (23) to obtain the node's stochastic service curve.

We illustrate the above calculations for Scenario 1 in Fig. 2. From Eq. (42) and (43), $\tau = 0.037$ and $\eta = 0.293$. Thus, $P_{nt} = 0.680$, $P_t = 0.320$ and $P_s = 0.027$. Again, we can see that the previous assumption that the first slot in $[s, s+t]$ is occupied by a transmission from the other $n-1$ node is a good approximation, as the transmissions by the other $n-1$ nodes happen much frequently than the given node when n is large (compare $P_t - P_s$ and P_s here) and also an *idle slot* is much smaller in length than L (39 idle slots here; in other words, the channel is often busy).

Fig. 5 shows I 's $\sigma(\theta), (\rho(\theta))$ -upper constrained characterization when θ ranges from 0.01 to 5.0.

For example, when $\theta = 0.1$, we have $\sigma_I(0.1) = 0.077$ and $\rho_I(0.1) = 0.924$. And

$$\begin{aligned}
\beta(t) &= (1 - r_I) \cdot t \\
g(x) &= \frac{e^{0.0077}}{1 - e^{0.0924 - 0.1r_I}} \cdot e^{-0.1x},
\end{aligned}$$

for any $1 > r_I > 0.924$.

Finally, we can not apply Proposition 1 to tightening $g(x)$, as $\rho(\theta) + \sigma(\theta) \geq 1$ here (considering $M_I(1) = e^\theta \leq e^{\rho(\theta) + \sigma(\theta)}$). In order to apply this proposition, we must

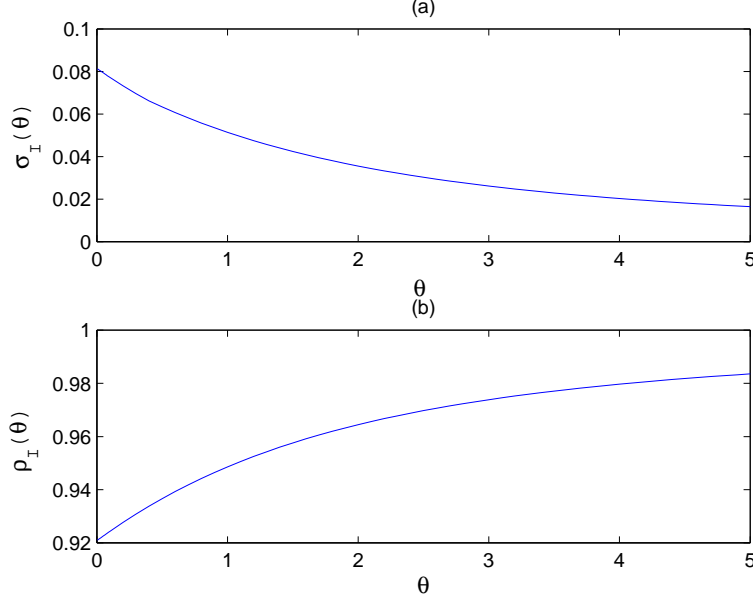


Fig. 5. The impairment process I 's $(\sigma(\theta), \rho(\theta))$ -upper constrained characterization

let $r_I \geq \rho(\theta) + \sigma(\theta) \geq 1$ which makes $\beta(t) = (1 - r_I)t \leq 0$.

4.3 Arrival Curves

In our performance evaluation we use Poisson traffic and we let λ be the average rate (packets/slot) of it. We have $a_A = \lambda$ by Definition 7, and

$$Ee^{\theta A(s,s+t)} = \sum_{i=0}^{\infty} \frac{(\lambda t)^i}{i!} e^{-\lambda t} e^{\theta i} = e^{\lambda t(e^\theta - 1)}. \quad (47)$$

. Therefore, Poisson traffic is $(\sigma_A(\theta), \rho_A(\theta))$ -upper constrained where $\sigma_A(\theta) = 0$ and $\rho_A = \frac{\lambda(e^\theta - 1)}{\theta}$.

We can get the vb arrival curves by Eq. (21). Note that we can improve the bounding function by $f(x) = e^{-\theta x}$ for $r_A \geq \rho_A(\theta)$ by Proposition 1 as Poisson process is time-independent.

As for the traffic in reality, we can get $Ee^{\theta A(s,s+t)}$ from traffic traces. Then we use Algorithm 1 to get the $(\sigma_A(\theta), \rho_A(\theta))$ -upper constrained characterization and the vb arrival curves.

4.4 Stability Condition and Backlog Bounds

By Proposition 4, the *stability condition of an 802.11 node in a WLAN* is

$$a_A < 1 - a_I = \frac{P_s \cdot L}{P_{nt} + P_t \cdot L}. \quad (48)$$

The stability condition in Scenario 1 is $a_A < 0.079$ packet/slot or 0.207Mbps by the 802.11 parameters in Fig. 3.

The backlog bounds is calculated immediately by Eq.(24) by plugging into traffic arrival curves and the 802.11 node's service curve. Note that it is an optimization problem depending on θ_1 and θ_2 . In general, we do not have an analytical solution for it. Since the problem dimension is very small, we can apply the method of exhaustion to get the optimal value.

We can also use Theorem 7 to improve on backlog bounds, as in our model we consider I under the saturated state which is independent of A . And it still needs to optimize the derived bounding function.

5 Performance Evaluation

In this section, we compare our backlog bounds derived in Section 4 with ns-2 simulations in Scenario 1 with Poisson traffic arrivals. The duration of each ns-2 simulation is 100 seconds which is long enough to let a node transmit thousands of packets. And we get the real $P\{B(t) > x\}$ over 100 independent simulations.

As shown in Section 4.4, we can derive stable backlog bounds when $a_A = \lambda < 0.079$ packet per slot. Fig. 6 plots the average backlog $E[B(t)]$ of ns-2 at $t = 50s$ and $\lambda = 0.077, 0.079$ and 0.081 packet/slot. We note that there is a sudden jump when $\lambda = 0.081$, indicating the critical point of stability is indeed around 0.079.

We use Jiang's model to calculate backlog bounds since the two models are proved to be equivalent in our studied case. There are two results of vb arrival curves for the traffic arrival process A : the *general* one (Theorem 5's Eq.(18)) and the *time-independent* one which improves on the former for the time-independent A (Proposition 1). And there are two relations between A and the impairment process I of the server: the *general* one (Theorem 2) and the *independent* one which improves on backlog bounds for independent A and I (Theorem 7).

So we can generate four results of backlog bounds.

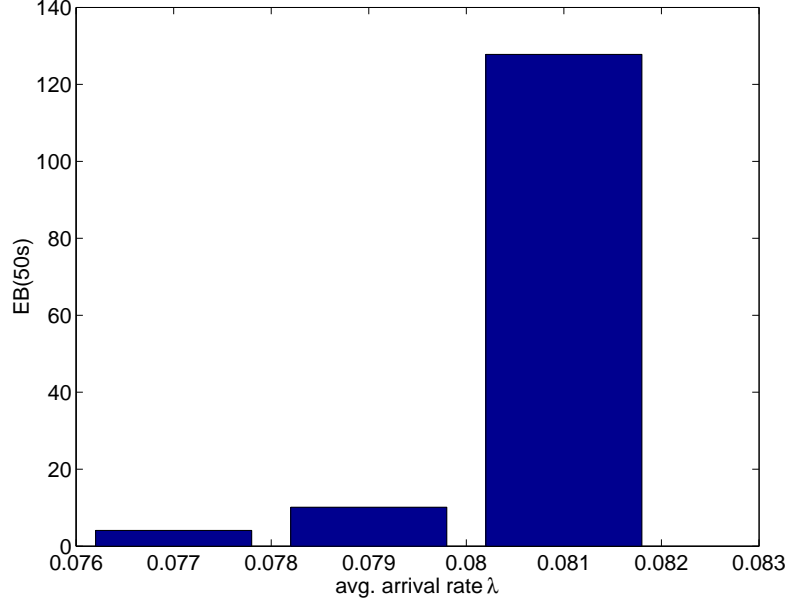


Fig. 6. $EB(t)$ ($t = 50s$) when $\lambda = 0.077, 0.079, 0.081$

- Bound 1: by *general* vb arrival curves of traffic and *general* backlog bounds.
- Bound 2: by *time-independent* vb arrival curves of traffic and *general* backlog bounds.
- Bound 3: by *general* vb arrival curves of traffic and *A-I-independent* backlog bounds.
- Bound 4: by *time-independent* vb arrival curves of traffic and *A-I-independent* backlog bounds.

Obviously, Bound 1 is the loosest and Bound 4 is the tightest among them.

To illustrate our results, we show the smallest x that makes $P\{B(t) > x\} \leq p$ for some probability p , i.e., $\min\{x : P\{B(t) > x\} \leq p\}$ where $p = 0.9, 0.8, \dots, 0.1, 0.05$. Obviously, for the same p , smaller x , tighter the bound. Fig. 7 (Fig. 8) shows our results when the 802.11 node's traffic arrival rate is $\lambda = 0.04$ (0.07) packet/slot.

We make the following observations. First, the backlog bounds improve significantly when we apply *time-independent* vb arrival curves for traffic or *A-I independent* case analysis. Second, the bounds of network calculus are much looser for higher traffic arrival rate while the real bounds of ns-2 simulations do not relax much. Note that A. Kumar's 802.11 model become very accurate near the saturated state [31] which is just the case here. The actual reason is: We have the constraint of $\rho_A(\theta_1) + \rho_I(\theta_2) < c$ (Eq.(24)). And we need to make θ_1 and θ_2 smaller to satisfy this constraint for higher traffic arrival rate, which leads to much looser bounding functions. Moreover, Theorem 5's Eq.(18) applies Boole's inequality to the bound functions of I , which are loose in general. Here brings the challenge for better network calculus models. Finally, we found in trace files that backlog bounds are sensitive to the parameters (i.e., θ_1 , θ_2 , r_A and r_I) and it is necessary to optimize

	$\min\{x : P\{B(t) > x\} \leq p\}$				
p	Bound 1	Bound 2	Bound 3	Bound 4	ns-2
0.9	24	8	7	7	1
0.8	25	9	8	8	1
0.7	25	10	8	8	1
0.6	25	10	9	9	1
0.5	26	11	10	10	1
0.4	27	12	10	10	1
0.3	28	13	11	11	1
0.2	29	14	12	12	2
0.1	31	17	14	14	2
0.05	33	19	16	16	4

Fig. 7. ns-2 and network calculus results of backlog bounds under $\lambda = 0.04\text{packet/slot}$ (low traffic load)

	$\min\{x : P\{B(t) > x\} \leq p\}$				
p	Bound 1	Bound 2	Bound 3	Bound 4	ns-2
0.9	201	61	50	50	1
0.8	203	64	55	55	1
0.7	206	68	58	58	1
0.6	209	72	63	63	2
0.5	212	76	66	66	2
0.4	217	82	71	71	2
0.3	223	89	77	77	4
0.2	231	99	85	85	4
0.1	245	114	98	98	6
0.05	258	129	109	109	7

Fig. 8. ns-2 and network calculus results of backlog bounds under $\lambda = 0.07\text{packet/slot}$ (severe traffic load)

them.

6 Related Work

In this section, we first present a brief overview of the theories and applications of stochastic network calculus and then the related works on the performance analysis of 802.11.

The increasing demand on transmitting multimedia and other real time applications over the Internet has motivated the study of quality of service guarantees. Towards it, deterministic and stochastic network calculus has been recognized by researchers as a promising step.

Essentially, the network calculus is the theory of queueing systems that comes from the seminal work by Cruz on the (σ, ρ) traffic characterization [2] [3] and work on the service curve characterization of Generalized Processor Sharing (GPS) schedulers [4] [5]. The theory has been developed by many researchers since then. The elegance of network calculus is due to the fundamental convolution formulas (under the min-plus algebra) that determine the departure process of a system from its arrivals and its service curve. The notable strength of the min-plus convolution is the ability to concatenate tandem nodes along a network path, and therefore network calculus has the ability to characterize the whole network as a single server, which is generally intractable by traditional queueing theory [1]. Le Boudec's book covers deterministic network calculus and its applications in the Internet [6]. Chang's book substantially presented the first approaches to stochastic network calculus besides deterministic network calculus [7]. Jiang summarized different types of stochastic arrival and service curves in a unified framework and proposed a new stochastic network calculus model stemmed from mb (maximal backlog centric) arrival curves, although its application conditions have some unsolved controversy. Jiang also wrote a book on the theory of stochastic network calculus [8]. Ciucu proposed an effective stochastic service curve that can be applied to concatenated systems and calculating end-to-end delay and backlog bounds, which exhibits a good scaling property of $O(H \log H)$ where H is the number of nodes traversed by a flow [17]. Ciucu also showed that his model can derive quite accurate delay bounds in M/M/1 and M/D/1 queueing systems by using the martingale technique [18]. More recently, Fidler proposed a novel solution of the queue system using expectations instead of probabilities [25], and he also made a comprehensive survey on the recent progress of stochastic network calculus [9]. Besides, Jiang wrote an overview on this topic from the queueing principle perspective and he presented a nice outlook by discussing many open challenges [10].

Many works have applied network calculus, for example, in measurement-based admission control schemes [19], in conformance testing, [20], in wireless sensor networks [21], in Aloha systems [26], in speeding up network simulations [22] [23], in bandwidth estimation [24] and even in manufacturing blocking systems in management science [27].

Compared with the existing theories of stochastic network calculus, we study the effectiveness of backlog bounds in a practical 802.11 WLAN by using the two classic stochastic network calculus models: Jiang’s and Ciucu’s. The latter can be extended to concatenated systems while the former still can not at the moment. Interestingly, we find that the two models can derive equivalent stochastic service curves and backlog bounds in our studied case, which can provide some hints for unifying the theories of stochastic network calculus in the future.

Existing works on the performance of 802.11 focus primarily on the throughput and capacity. Bianchi proposed a Markov chain model of 802.11 [29]. A. Kumar et al. proposed a probability model of 802.11 [30] which simplifies Bianchi’s model and it is shown to be quite accurate even in the multi-hop case [31]. In our paper, we adopt A. Kumar’s model to derive the stochastic service curves of an 802.11 node. There are some works on 802.11 queueing analysis based on traditional queueing theory. Zhai et al. assumed Poisson traffic arrival and proposed an M/G/1 queueing model of 802.11 [32]. Tickoo proposed a G/G/1 queueing model of 802.11 [33] [34]. Bredel and Fidler modeled the 802.11 DCF as a fluid GPS scheduler yielding a fair average service rate [25]. And Ciucu analyzed the non-asymptotic throughput and delay distribution in multi-hop wireless networks by network calculus approach considering Aloha systems [26].

Compared to existing analysis of 802.11, we are the first to analyze the *concrete* 802.11 transmissions by Jiang’s and Ciucu’s models and study their *effectiveness* on bounding backlogs. We show that stochastic network calculus is effective *theoretically* in that the bounds imply stable backlogs as long as the average arrival rate is less than the average service rate. However, the bounds are quite loose and we show that they can be improved significantly for time-independent arrivals or under the independent cases of arrival and service processes. And we note that it is still a challenge for a better theory of stochastic network calculus towards tighter bounds in practice. Therefore, our work offers a good reference to applying stochastic network calculus to practical scenarios.

7 Conclusion and Future Work

In this paper, we present concrete computations of 802.11 backlog bounds and study the bounds effectiveness using stochastic network calculus, from general models to detailed calculations. We model a wireless node as a single server with impairment service based on two best-known models in stochastic network calculus: Jiang’s [16] and Ciucu’s [17]. And we find that they can derive equivalent stochastic service curves and backlog bounds in our studied case. Then we care about the effectiveness of network-calculus backlog bounds *theoretically*. And we prove that the network-calculus backlog bounds imply stable backlog as long as the average rate of traffic arrival is less than that of service. Next, we consider

the effectiveness of network-calculus bounds in practice. We derive the stochastic service curve of an 802.11 node from A. Kumar's 802.11 model, which is crucial to get backlog bounds. We observe the derived bounds are loose when compared with ns-2 simulations. However, the martingale and independent case analysis techniques can improve the bounds significantly. But still the bounds are not tight. We note the reason is due to the looseness in network calculus itself such as Theorem 5's Eq.(18). The open questions are: Can we find tighter bounding functions under certain conditions? How do we optimize on stochastic arrival/service curves? Furthermore, does there exist any unified theory of all network calculus models? And these are the future works.

8 Acknowledgement

We thank Prof. Yuming Jiang with NTNU, Prof. Florin Ciucu with TU-Berlin, and Dr. Kai Wang with Tsinghua and Caltech for their nice discussions on the theories of stochastic network calculus. We thank Prof. Jianming Zhu and Prof. Haiqi Feng in our school for holding related seminars.

References

- [1] L. Kleinrock, "Queueing Systems. Volume 1: Theory," *Wiley Interscience*, 1975.
- [2] R. L. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation," *IEEE Trans. Information Theory*, vol. 37(1), pp. 114-131, 1991.
- [3] R. L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis," *IEEE Trans. Information Theory*, vol. 37(1), pp. 132-141, 1991.
- [4] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-node Case," *IEEE/ACM Trans. Networking*, vol. 1(3), pp. 344-357, 1993.
- [5] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple-node Case," *IEEE/ACM Trans. Networking*, vol. 2(2), pp. 137-150, 1994.
- [6] J.-Y. Le Boudec and P. Thiran, "Network Calculus: A Theory of Deterministic Queueing Systems for the Internet," *Springer-Verlag*, 2001.
- [7] C.-S. Chang, "Performance Guarantees in Communication Networks," *Springer-Verlag*, 2000.
- [8] Y. Jiang and Y. Liu, "Stochastic Network Calculus," *Springer*, 2008.

- [9] M. Fidler, "A Survey of Deterministic and Stochastic Service Curve Models in the Network Calculus," *IEEE Communications Surveys & Tutorials*, vol. 12(1), 2010.
- [10] Y. Jiang, "Stochastic Network Calculus for Performance Analysis of Internet Networks C An Overview and Outlook," Proc. International Conference on Computing, Networking and Communications (ICNC), 2012.
- [11] R. L. Cruz, "Quality of Service Guarantees in Virtual Circuit Switched Networks," *IEEE J. Selected Areas Communication*, vol. 13(6), pp. 1048-1056, 1995.
- [12] C.-S. Chang, "On Deterministic Traffic Regulation and Service Guarantees: A systematic Approach by Filtering," *IEEE Trans. Information Theory*, vol. 44(3), pp. 1097-1110, 1998.
- [13] C. Li, A. Burchard, and J. Liebeherr, "A Network Calculus with Effective Bandwidth," Technical Report, CS-2003-20, University of Virginia, 2003; *IEEE/ACM Trans. Networking*, vol. 15(6), pp. 1442-1453, 2007.
- [14] Y. Jiang and P. J. Emstad, "Analysis of Stochastic Service Guarantees in Communication Networks: A Traffic Model," Proc. *International Teletraffic Congress*, 2005.
- [15] Y. Jiang and P. J. Emstad, "Analysis of Stochastic Service Guarantees in Communication Networks: A Server Model," Proc. *IEEE IWQoS*, 2005.
- [16] Y. Jiang, "A Basic Stochastic Network Calculus," Proc. *ACM SIGCOMM*, pp. 123-134, 2006.
- [17] F. Ciucu, A. Burchard, and J. Liebeherr, "A Network Service Curve Approach for the Stochastic Analysis of Networks," Proc. *ACM SIGMETRICS*, vol. 33(1), 2005.
- [18] F. Ciucu, "Network Calculus Delay Bounds in Queueing Networks with Exact Solutions," Proc. *International Teletraffic Congress*, 2007.
- [19] Y. Jiang, P. Emstad, A. Nevin, V. Nicola, and M. Fidler, "Measurement-Based Admission Control for a Flow-Aware Network," Proc. *EuroNGI 1st Conference on Next Generation Internet Networks - Traffic Engineering*, 2005.
- [20] Y. Liu, C.-K. Tham, and Y. Jiang, "Comformance Analysis in Networks with Service Level Agreements," *Computer Networks*, vol. 47, pp. 885-906, 2005.
- [21] J. B. Schmitt and U. Roedig, "Sensor Network Calculus - A Framework for Worst Case Analysis," Proc. *DCOSS*, pp. 141-154, 2005.
- [22] H. Kim and J.C. Hou, "Network Calculus Based Simulation for TCP Congestion Control: Theorems, Implementation, and Evaluation," Proc. *IEEE INFOCOM*, pp. 2844-2855, 2004.
- [23] H. Kim, "Integrating Network-Calculus-Based Simulation with Packet-Level Simulation for TCP-Operated Networks," *Computer Networks*, vol. 50(12), pp. 1995-2012, 2006.
- [24] J. Liebeherr, M. Fidler, and S. Valaee, "A Min-Plus System Interpretation of Bandwidth Estimation," Proc. *IEEE INFOCOM*, pp. 1127-1135, 2007.

- [25] M. Bredel and M. Fidler, "Understanding Fairness and Its Impact on Quality of Service in IEEE 802.11," *Proc. IEEE INFOCOM*, pp. 1098-1106, 2009.
- [26] F. Ciucu, O. Hohlfeld, and P. Hui, "Non-Asymptotic Throughput and Delay Distributions in Multi-Hop Wireless Networks," *Proc. Allerton*, 2010.
- [27] A. Bose, X. Jiang, B. Liu, and G. Li, "Analysis of Manufacturing Blocking Systems with Network Calculus," *Performance Evaluation*, vol. 63(12), pp. 1216C1234, 2006.
- [28] IEEE Std. 802.11, "Wireless LAN Media Access Control (MAC) and Physical Layer (PHY) Specifications," <http://standards.ieee.org/getieee802/>, 1999.
- [29] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE J. Selected Area in Communication*, 18(3): 535-547, 2000.
- [30] A. Kumar, E. Altman, D. Miorandi, and M. Goyal, "New Insights from a Fixed-Point Analysis of Single Cell IEEE 802.11 WLANs," *IEEE Trans. Networking*, pp. 588-601, 2007.
- [31] Y. Gao, J.C.S Lui, D.-M. Chiu, "Determining the End-to-end Throughput Capacity in Multi-hop Networks: Methodology and Applications," *Proc. ACM SIGMETRICS/Performance*, pp. 39-50, 2006.
- [32] H. Zhai and Y. Fang, "Performance of Wireless LANs Based on IEEE 802.11 MAC Protocols," *Proc. IEEE PIMRC*, vol. 3, pp. 2586-2590, 2003.
- [33] O. Tickoo and B. Sikdar, "Queueing Analysis and Delay Mitigation in IEEE 802.11 random access MAC based Wireless networks," *Proc. IEEE INFOCOM*, vol. 2, pp. 1404-1413, 2004.
- [34] O. Tickoo and B. Sikdar, "A queueing model for finite load IEEE 802.11 random access MAC," *Proc. IEEE ICC*, vol. 1, pp. 175-179, 2004.

Appendix A: Proof of Proposition 1

Proof:

For a fixed t , we construct a stochastic process $X(s) = e^{\theta(A(t-s,t)-rs)}$ ($0 \leq s \leq t$) and we have $X(s+1) = X(s)e^{\theta(A(t-s-1,t-s)-r)}$. We will show that if $a(s) \equiv A(s) - A(s-1)$ is independent for each time slot s and $r \geq \rho(\theta) + \sigma(\theta)$, then $X(s)$ is supermartingale, i.e., $E[X(s+1)|X(0), \dots, X(s)] \leq X(s)$.

Because $a(s)$ is time-independent, we have

$$\begin{aligned} E[X(s+1)|X(0), \dots, X(s)] &= X(s) \cdot E[e^{\theta(a(t-s)-r)}] \\ &= X(s) \cdot e^{-\theta r} \cdot Ee^{\theta a(t-s)}. \end{aligned} \tag{49}$$

Because $A(t)$ is $(\sigma(\theta), \rho(\theta))$ -upper constrained, we have $Ee^{\theta a(s)} \leq e^{\rho(\theta) + \sigma(\theta)}$ for all $s \geq 0$. When $r \geq \rho(\theta) + \sigma(\theta)$ and by Eq. (49), we have

$$\mathbb{E}[X(s+1)|X(0), \dots, X(s)] \leq X(s). \quad (50)$$

Thus, $X(s)$ is a supermartingale.

Doob's martingale inequality says that $P\{\sup_{0 \leq s \leq t} X(s) \geq k\} \leq \frac{\mathbb{E}X(0)}{k}$ when $X(s)$ is a supermartingale (note: $\mathbb{E}X(0) = 1$ here) for any constant k . Let $k = e^x$, we have

$$\begin{aligned} & P\left\{\sup_{0 \leq s \leq t} [(A(s, t) - r(t - s))] > x\right\} \\ &= P\left\{\sup_{0 \leq s \leq t} [X(s)] \geq e^x\right\} \leq e^{-x}. \end{aligned} \quad (51)$$

■

Appendix B: Algorithm 1 (*Numerical Calculation of $\sigma_I(\theta)$ and $\rho_I(\theta)$*)

Let $y(t) = \sup_{s \geq 0} \{\frac{1}{\theta} \log \mathbb{E}e^{\theta I(s, s+t)}\}$. Obviously, $y(t)$ is an increasing function of t with $y(0) = 0$. We define axes t and axes t_\perp (vertical to t) on a plane and we can imagine plotting $y(t)$ on it. We define the slope of $y(t)$, $s(t) = y(t) - y(t-1)$.

We calculate $s(t)$ for $t = 1, 2, 3, \dots$ until it converges at some t^* , i.e., $(1 - \epsilon) \cdot s(t^* - 1) \leq s(t^*) \leq (1 + \epsilon) \cdot s(t^* - 1)$ where ϵ is a small number, e.g. 10^{-5} .

We draw a straight line $l(t)$ with the slope $s(t^*)$ crossing the point $(t^*, y(t^*))$ on the axes of t and t_\perp . Obviously, the line crosses the point $(0, y(t^*) - s(t^*)t^*)$. The maximum vertical distance between $y(t)$ and $l(t)$, $v_m = \max_{0 \leq t \leq t^*} \{y(t) - l(t)\}$. We shift $l(t)$ by v_m and get $\tilde{l}(t)$. Clearly, $y(t) \leq \tilde{l}(t)$. By Definition 6, $\rho(\theta) = s(t^*)$ and $\sigma(\theta) = y(t^*) - s(t^*)t^* + v_m$.

Yue Wang's Biography

Yue Wang received his B.Sc degree in Mechanics and M.Sc degree in Computer Science from Peking University, and received his PhD degree in The Chinese University of Hong Kong. He is currently a lecturer with School of Information, Central University of Finance and Economics in China. His research interests lie in the performance evaluation and optimization of wireless networks. He has served on the TPC of Workshop on Network Calculus (WoNeCa) in 2012. And he is a member of ACM and CCF (China Computer Federation).